Goldspink C. & Kay R, 2009, 'Agent Cognitive Capability and Orders of Emergence', forthcoming in Trajkovski, Goran and Collins, Samuel (eds), "Agent-Based Societies: Social and Cultural Interactions"

Agent Cognitive Capabilities and Orders of Social Emergence.

Christopher Goldspink
Centre for Research in Social Simulation
Department of Sociology
University of Surrey. Guildford, GU1 7XH, UK
c.goldspink@surrey.ac.uk

Robert Kay
Co-founder, Incept labs
432 Kent St Sydney 2000
rkay@inceptlabs.com.au

Dr. Chris Goldspink has twenty years experience as an internal and external consultant in organisation and management improvement. His background is diverse and includes experience with information technology, education, organizational management and public sector reform. He has senior and middle management experience and maintains an interest in practical management in both the public and private sectors, teaching and consulting in both. He teaches in areas including strategic management, leadership and marketing as well as public management reform, the Australian system of Government and in systems derived research methodologies. His research interests include the applicability of complex systems theory to social and organizational analysis, computer simulation of social phenomena and issues in cross cultural education.

Dr Robert Kay is co founder of Incept Labs, a firm specializing in strategic innovation and commercialization services. Formerly Robert was the Head of Strategic Innovation at Westpac Banking Corporation. This role involved the integration of various systems theories into the innovation processes of the bank. He has also performed as a Senior Lecturer in Information Systems at the University of Technology, Sydney (UTS), and lectured at the University of Western Sydney; the International Graduate School of Management, University of South Australia; and University of New South Wales. His central research interest is in the application of autopoietic and complexity theory to understanding organisational sustainability and change.

Agent Cognitive Capabilities and Orders of Social Emergence.

ABSTRACT

This chapter critically examines our theoretical understanding of the dialectical relationship between emergent social structures and agent behaviors. While much has been written about emergence individually as a concept and the use of simulation methods are being increasingly applied to the exploration of social behavior, the concept of 'social emergence' remains ill defined. Furthermore there has been little theoretical treatment or practical exploration of how both the range and type of emergent structures observed may change as agents are endowed with increasingly sophisticated cognitive abilities. While we are still a very long way from being able to build artificial agents with human-like cognitive capabilities it would be timely to revisit the extent of the challenge and to see where recent advances in our understanding of higher order cognition leave us. This chapter provides a brief recount of the theory of emergence, considers recent contributions to thinking about orders of emergence, and unpacks these in terms of implied agent characteristics. Observations are made about the implications of alternative cognitive paradigms and the position is proposed that an enactivist view provides the most logical pathway to advancing our understanding. The chapter concludes by presenting an account of reflexive and non-reflexive modes of emergence which incorporates this view.

INTRODUCTION

Building and working with artificial societies using the methods of multi-agent social simulation serves us in several ways: 1) It allows us to operationalize social theories and to compare simulated behaviors with those observed in the real world; and 2) it allows us to build new theory by exploring the minimal mechanisms that might explain observed social behavior. Most importantly 3) it provides a unique ability to explore the interplay between levels of phenomena and to understand dynamic properties of systems. A great deal can and has been achieved in both these areas with even the simple methods we currently have available. However, Keith Sawyer (2003) has recently reminded us that, to date, we have worked with agents with very limited cognitive capability and that this necessarily limits the range and type of behavior which can be explored. This echoes a sentiment made a decade ago by Christiano Castelfranchi (1998a) that social simulation is not really *social* until it can provide an adequate account of the implication of feedback between macro and micro which becomes possible with higher cognitive functioning of social agents.

In many respects, developments in our capacity to simulate artificial societies have led us to confront anew a long-standing issue within social theory. This is a problem that social science conducted within traditional disciplinary boundaries has become quite adept at avoiding. Indeed it can be argued that the particular form disciplinary fragmentation takes in social science is a primary strategy for avoiding it. The problem is referred to in a number of ways depending on the disciplinary tradition. This chapter begins by revisiting this most important of problems. In terms of the challenge it poses to artificial societies it can be expressed in the following three questions.

1. What are the fundamental cognitive characteristics which distinguish human agents from animal or automaton?

2. How do these characteristics influence the range and type of behaviors agents may generate and the emergent structures which they may give rise to?
3. How can we theorize about the relationship between cognitive capability and categories of emergent form?

These questions form the focus for this chapter. We begin to address them by revisiting the contribution of alternative schools of thought to our understanding of the nature and origins of emergent structure and alternative concepts of orders of emergence. We then discuss the implications of the two competing cognitive paradigms within AI – that of cognitivism and the enactive view. Finally we turn to current research on the development of human cognition and examine its implications for anticipating different orders of emergent structure – proposing what we call reflexive and non-reflexive classes of emergence. Finally a research program for the advancement of understanding in this area is proposed.

This work has its origins in two strands of research with which the authors are currently involved. The first addresses the relationship between micro and macro levels of social behavior and organization directly. Over the past decade we have explored the characteristics of the micro-macro problem (see Chris Goldspink & Kay, 2003, 2004) in pursuit of a coherent and consistent account of the interpenetration (circular causality) between micro and macro phenomena. Our aim is to develop a theory which can provide a substantive account of fundamental social generative mechanisms. To date no such social theory exists that satisfactorily explains this dynamic.

The other strand is one author's involvement with the Centre for Research in Social Simulation and though it the European Union funded project titled Emergence in the Loop (EMIL). The aim of EMIL is to : a) provide a theoretical account of the mechanisms of normative self-regulation in a number of computer mediated communities b) specify the minimum cognitive processes agents require to behave in normative ways c) develop a simulator which can replicate the range and type of normative behavior identified by the empirical research so as to further deepen our understanding of how and under what conditions normative self-regulation is possible and the range and type of environmental factors which influence it.

A BRIEF RECOUNT OF THE THEORY OF EMERGENCE

The notion of emergence has a long history, having been invoked in a number of disciplines with varying degrees of centrality to the theoretical and methodological development of associated fields. Unfortunately the concept has largely remained opaque and ambiguous in its conceptualization, leading to the criticism that it stands as little more than a covering concept – used when no adequate account or explanation exists for some unexpected phenomena. Clayton has argued that the concept covers:

> *…a wide spectrum of ontological commitments. According to some the emergents are no more than patterns, with no causal powers of their own; for others they are substances in their own right… (Clayton, 2006: 14).*

The origin of the concept has been attributed to George Henry Lewes who coined the term in 1875 (Ablowitz, 1939). It subsequently found wide adoption within the philosophy of science but more recently has been advanced within three distinct streams: *philosophy*, particularly of science and mind; *systems theory*, in particular complex systems; and *social science* where it has largely been referred to under the heading of the micro-macro link and/or the problem of structure and agency. Interestingly there has been relatively little cross fertilization of thinking between these streams.

The Contribution from Philosophy of science

The philosophy of science and philosophy of mind stream is arguably the oldest – some date it back to Plato (Peterson, 2006) but the debate is widely seen as having come to focus with the British Emergentists (Eronen, 2004; Shrader, 2005; Stanford Encyclopaedia of Philosophy, 2006). This school sought to deal with the apparent qualitatively distinct properties associated with different phenomena (physical, chemical, biological, mental) in the context of the debate between mechanism and vitalism: the former being committed to Laplacian causal determinism and hence reductionism and the latter invoking 'non-physical' elements in order to explain the qualitative difference between organic and in-organic matter. This stream remains focused on explaining different properties of classes of natural phenomena and with the relationship between brains and minds (See Clayton & Davies, 2006 for a recent summary of the positions). As a consequence this has been the dominant stream within artificial intelligence. Peterson (2006: 695) summarizes the widely agreed characteristics of emergent phenomena within this stream as follows. Emergent entities:

1. Are characterized by higher-order descriptions (i.e. form a *hierarchy*).
2. Obey higher order *laws*.
3. Are characterized by *unpredictable novelty*.
4. Are *composed of* lower level entities, but lower level entities are *insufficient* to fully account for emergent entities (*irreducibility*).
5. May be capable of *top-down causation*.
6. Are characterized by *multiple realization or wild disjunction* (Fodor, 1974) (alternative micro-states may generate the same macro states).

A key concept within these discussions is that of *supervenience*: a specification of the 'loose' determinisms held to apply between levels such that *'…an entity cannot change at a higher level without also changing at a lower level'* (Sawyer, 2001: 556). Within this stream prominence of place is given to both downward and upward causation. Clayton and Davies (2006) specify downward causation as involving macro structures placing *constraint* on lower level processes hence *'Emergent entities provide the context in which local, bottom up causation takes place and is made possible'* (Peterson, 2006: 697). Davies (2006) argues that the mechanism of downward causation can usefully be considered in terms of boundaries. Novelty, he argues, may have its origin in a system being 'open'. If novel order emerges it must do so within the constraints of physics. He concludes:

> …top-down talk refers not to vitalistic augmentation of known forces, but rather to the system harnessing existing forces for its own ends. The problem is to understand how this harnessing happens, not at the level of individual intermolecular interactions, but overall – as a coherent project. It appears that once a system is sufficiently complex, then new top down rules of causation emerge. (Davies 2006: 48).

For Davies then, top-down causation is associated with self-organization and may undergo qualitative transitions in form with increasing system complexity. For Davies also it is the 'openness' of some systems that 'provides room' for self-organizing process to arise, but he concludes, *'openness to the environment merely explains why there may be room for top-down causation; it tells us nothing about how that causation works.'* The devil then, is in the detail of the mechanisms specific to particular processes in particular contexts and particular phenomenal domains. Perhaps then a part of the problem with the concept is that it has been approached at too abstract a level.

The contribution from Social Science

The micro-macro problem – the relationship between the actions of individuals and resulting social structures and the reciprocal constraint those structures place on individual agency –

has long standing in social science as well as in philosophy. The problem is central to many social theories developed throughout the 19th and 20th century. Examples include: Marxian dialectical materialism (Engels, 1934) built upon by, among others, Vygotsky (1962) and Lyont'ev (1978); the social constructionism of Berger and Luckmann (1972); Gidden's structuration theory (1984); and the recent work of critical realists (Archer, 1998; Archer, Bhaskar, Ciollier, Lawson, & Norrie, 1998; Bhaskar, 1997, 1998). These alternative theories are frequently founded on differing assumptions, extending from the essentially objectivist/rationalist theory of Coleman (1994), through the critical theories of Habermas and to the radical constructivism of Luhmann (1990; 1995).

Fuchs & Hofkirchner (2005: 33) have recently suggested a four category schema for classifying social theory according to the ontological position adopted with respect to the micro-macro relationship. The majority of existing social theories, they argue, fall into one or other of two categories which they label *individualism* and *sociologism*. Neither of these 'paradigms' provides a theoretical foundation which supports exploration let alone the possibility of advancing understanding of the interplay between agency and structure, rather the problem is avoided by restricting analysis to one level or the other. A third category, *dualism*, while considering both aspects, insists on the adoption of a dichotomous stance and as a consequence does not support any understanding of the interplay between levels. Only those theories categorized as *dialectical* therefore have relevance. Even here, it is reasonable to conclude that little practical advance has been achieved, as most positions result in a straddling of bottom up and top-down arguments and/or suffer from excessively vague conceptualization. These theories also quickly break down into a dichotomy the moment an attempt is made to make them operational.

What has been largely agreed, despite the very different theoretical and often inadequate handling of this problem, is that structure and agency come together in *activity* or in *body-hood* – the specific psycho-motor state at the instant of enaction. Both Vygotsky and Giddens, for example, focus on action as the point of intersection between human agency and social structures and it is implicit in Bourdieu's *habitus* also.

The contribution from Systems Theory

Systems language was clearly evident in the work of the early Emergentists and in a great deal of sociology and anthropology which took seriously the structure/agency problem – notably that of Margaret Mead and Gregory Bateson. However, 'systems' as a focus of systematic research arguably took form with von Bertalanffy's attempt to establish a General Systems Theory in 1950 (Bertalanffy, 1950; Bertalanffy_von, 1968). As the science of 'wholes' systems theory stands in contrast to reductionisms concern with parts. Systems theory was put forward as a counter to what was perceived as excessive reductionism dominating scientific discourse during much of the 20th century.

In the early stages of development of the theory systems tended to be modeled as 'black boxes' effectively masking the relationship between micro and macro elements. The application of the concept to social science, in particular through the development by Ernst von Glasersfeld and Heinz von Foerster (Keeney, 1987) of social cybernetics along with soft systems approaches (Checkland, 1988) provided a theoretical lens and methods useful for describing the systemic behavior of social systems. So while the aspiration of GSM to establish a general science of systems is generally regarded to have failed (Jackson, 2000), systems approaches have contributed valuable methods for the study of the interplay between levels in a social system. The Systems view of emergence was founded on:

- Holism; the whole is greater than the sum of its parts.

- A concern with *feedback both positive and negative*.
- A concern with boundaries and boundary conditions.

More recently the development of complex systems theory and its application to natural, social and cognitive phenomena has provided additional concepts upon which much current debate about emergence draws. Many of these concepts and methods have become widely used within the multi-agent modeling community (Castelfranchi, 1998b; Conte, Hegselmann, & Terna, 1997; Gilbert, 1995; Holland, 1998).

Within contemporary debate, and in contrast to the position taken by the British Emergentists who argued that irreducibility was the *exception* (Eronen, 2004), most real world systems are now argued to be non-linear (S. Kauffman, 2000; S. A. Kauffman, 1993, 1996; Stewart, 1990) and hence irreducible. It is non-linearity which contributes to these system's capacity for novelty and unpredictability through the presence of deterministic Chaos (Lorenz, 2001; Williams, 1997) and/or equifinality. Equifinality as it is known within systems theory, or the principle of 'wild disjunction' as it is known in philosophy, refers to a system where a single high level property may be realized by more than one set of micro-states which have no lawful relationship between them (Richardson, 2002a, 2002b; Sawyer, 2001). As there is no a-priori basis by which the likely micro state can be determined, such systems are irreducible and unpredictable in principle.

Observations

The concept of emergence has led to the establishment of a number of general principles which describe the relationship between micro and macro phenomena, as well as some methods and techniques for identifying and exploring it. Specifically, we can conclude that there are systems which are:

- inherently analytically reducible (to which the concept of emergence does not apply);
- analytically reducible in principle but difficult to reduce in practice and/or where an advance in science/knowledge is needed for reduction to be possible because the results were 'unexpected' (Chalmers, 2006) (to which the concept of 'weak' emergence can be applied);
- not reducible in principle (to which the principle of 'strong' emergence is relevant).

We argue that all living systems and all social systems belong to the latter class. Accordingly we agree with McKelvey (1997) that a great deal of social order may be attributable to complex organization involving non-liner relations between elements. It is for this reason that simulation methods are regarded as important but only to the extent that we can construct artificial societies which are reasonable analogues of the social systems we want to understand and this implies agent architectures which are capable of generating the range of social behaviors/structures of interest. The problem here is that we still have a very rudimentary understanding of what cognitive capabilities support or are necessary for what range and types of social structures.

In the following section we draw on the limited prior attention given to this problem and attempt to clarify what is currently known. Throughout the discussion, pointers are provided to where the mechanisms being outlined have, at least in part, been incorporated into computer simulations of artificial intelligence or artificial societies.

ORDERS OF EMERGENCE

A number of authors have identified what they refer to as orders of emergence. Gilbert, for example distinguishes between first and second order emergence. First order emergence includes macro structures which arise from local interactions between agents of limited cognitive range (particles, fluids, reflex action). By contrast, second order emergence is argued to arise *'where agents recognise emergent phenomena, such as societies, clubs, formal organizations, institutions, localities and so on where the fact that you are a member or a non-member, changes the rules of interaction between you and other agents.'* (Gilbert, 2002). This reflects high order cognition on the part of the agent. In particular it reflects a range of capabilities including but not limited to the ability to distinguish class characteristics; assess 'self' for conformity with class characteristics and/or signals from other agents which suggest acceptance or belonging; the ability to change rule associations and behavior as a function of these changes. First and second order emergence then each imply qualitatively distinct mechanisms and suggest a continuum of orders of emergence linked, in biological entities at least, to cognitive capability.

In a similar vein, Castelfranchi (1998a: 27) has distinguished what he refers to as cognitive emergence. *'Cognitive emergence occurs where agents become aware, through a given 'conceptualization' of a certain 'objective' pre-cognitive (unknown and non deliberated) phenomenon that is influencing their results and outcomes, and then, indirectly, their actions.'* This approach is based on a first generation AI (Franklin, 1998) approach to conceptualizing agents: agent cognition is assumed to involve acting on beliefs desires and intentions (BDI). Thus Castelfranchi conceives of a feedback path from macro pattern to micro behavior in much the same way as Gilbert, except that here a cognitive mechanism is specified. Castelfranchi argues that this mechanism has a significant effect on emergence and indeed *'characterises the theory of social dynamics'* – that is, it gives rise to a distinct class of emergent phenomena. In this account, the representations agents have about the beliefs, desires and intentions of other agents plays a causal role in their subsequent behavior and therefore shapes the structures they participate in generating. In this same chapter Castelfranchi argues that understanding this process is fundamental to social simulation: it is where social simulation can make its greatest contribution.

These ideas are more comprehensively reflected in the five orders of emergence suggested by Ellis (2006:99-101). These are:

1. Bottom up leading to higher level generic properties (examples include the properties of gases, liquids and solids)
2. Bottom up action plus boundary conditions leading to higher level structures (e.g. convection cells, sand piles, cellular automata)
3. Bottom up action leading to feedback and control at various levels leading to meaningful top down action - teleonomy (e.g. living cells, multi-cellular organisms with 'instinctive' – phylogenetically determined reactive capability)
4. as per 3 but with the addition of explicit goals related to memory, influence by specific events in the individuals history (i.e. learning)
5. In addition to 4 some goals are explicitly expressed in language (humans).

Ellis's framework makes clear that the range and type of emergence possible in a system depends fundamentally on the range and class of behavior agents are able to generate and that this varies depending on the properties of the agent.

If we consider Ellis' category one emergence, it is apparent that particles have fixed properties and are able to enter into a limited range of interactions (specified by physical laws) based on those properties. Swarms of particles can nevertheless demonstrate some rudimentary self-organization and hence emergence (Kennedy & Eberhart, 2001). Physics has furnished good accounts of many specific examples (Gell-Mann, 1995) but they have limited implication for our understanding of social behavior.

Category two has also recently been well explored as it is the focus of complexity theorists. Examples include the work of Per Bak (1996) on sand piles and earthquakes, Lorenz (2001) on weather systems and Prigogine (1997; 1985) on far from equilibrium systems. Many so called social simulations also belonging here– specifically those which incorporate agents which have fixed behaviors and no capacity for learning (individual or social). These include classic simulations based on swarms (Boids) and/or involving fixed decision criteria or rules such Schelling's segregation model, the cooperation models of Axelrod (1984) or the Sugarscape models of Epstein and Axtell (1996). Some may argue that these models involve agents with goals and therefore represent examples of fourth order emergence. The transition between 3rd order and fourth, as will be argued below, involves a move to agent autonomy that is missing in these models: their goals are designed in and not a result of their own operation it is for this reason that we argue they belong to order two.

It is significant that Ellis provides primarily biological examples for his category three order of emergence. The paradigmatic biological entity which illustrates the processes of reciprocal micro-macro causality and for which we have an excellent description which has been made operational both in vitro and in silico (see for example McMullin & Grob, 2001; F. Varela, Maturana, & Uribe, 1974) is the cell. While the mechanisms of autocatalysis and the metabolic pathways of cell self-production are well known, well documented and closely studied, the most concise articulation of the fundamental processes involved come with the theory of autopoiesis developed by the theoretical biologists Humberto Maturana and Francisco Varela (H. Maturana & Varela, 1980; H. R. Maturana & Varela, 1992; F. Varela, 1979; F. Varela et al., 1974). Unfortunately this account is not widely appreciated even within biology itself[1]. Varela (1997: 78) states:

> Autopoiesis is a prime example of a …dialectics between the local component levels and the global whole, linked together in reciprocal relation through the requirement of constitution of an entity that self-separates from its background.

The theory of autopoiesis provides a foundation for understanding other emergent processes, particularly those associated with biological entities. The originating authors themselves extended it to cover multi-cellular entities and to provide a more general theory of cognition. Others have gone so far as to argue that it furnishes a theory of society and/or organization (Niklas Luhmann, 1995; von_Krogh & Roos, 1995; Zeleny, 1991) although this remains controversial (Bednarz, 1988; Mingers, 2002, 2004) and we specifically reject it as incompatible with the original concept and as unnecessary (C. Goldspink, 2000; Kay, 1999).

Unlike the self-organizing processes which characterize the second order, the defining characteristic of biological self-organization is the attainment of 'strong autonomy' (Rocha, 1998). While Ellis does not say so directly, it would appear that it is the advent of a self-referential operational closure which demarcates third and higher orders of emergence from the lower orders.

[1] Quite why this should be the case is not clear. It does challenge the dominant paradigm within molecular biology and may have been displaced by the apparent potential offered by genomics (Oyama, 2000). It may also be that its implications are most significant outside of the biology discipline.

Maturana and Varela argue that cognition is associated with this operational closure or autonomy. Autonomy is used here to refer to a *constitutive* process rather than as a *categorical* distinction and cognition is defined as the range of behaviors the agents can generate to remain *viable* or *to retain its identity* as a self-constituting agent (Froesea, Virgo, & Izquierdo, 2007; Thompson & Varela, 2001). For those immersed in symbolic AI it may come as a surprise that a biological cell may thus be described as a cognitive entity. This theme will be developed further in a following section as it is central to the idea of enactive cognition finding increasing uptake within second generation AI, artificial life and robotics (Barandiaran, 2005; Di Paolo & Lizuka, 2007; Di Paolo, Rohde, & De Jaegher, 2007; Moreno & Etxeberria, 1995; Moreno, Umerez, & Ibanes, 1997).

In his third order category Ellis includes a range of capabilities of biological entities up to and including 'instinctive' action. These suggest that this category would pertain to single and multi-cellular organisms including those with a central nervous system. It may be that this order is too broadly cast. Multi-cellularity is arguably another threshold point as differentiated aggregates of cells display greater capacity to respond to their environment, even where they do not possess a central nervous system, than do individual cells. Furthermore those with a central nervous system enjoy even greater behavioral flexibility. As a consequence each probably originates a distinct macro phenomenology different from that of the cells that constitute them (H. R. Maturana & Varela, 1992).

The primary point of distinction between order three and order four would appear to be between (phylogenetically) fixed individual characteristics and a capacity for an individual agent to have goals and to learn. The mechanisms by which these characteristics are acquired and fixed at the level of individuals (sexual transmission and natural selection) are ignored by Ellis or seen as unimportant from the perspective of emergence. This is reasonable if our concern is with social behavior which manifests over relatively short time cycles in geological terms. When does a capacity to adjust structure in response to an environment as implied by the characteristics of Ellis' third order become the learning ability associated with the forth order?

Ellis explicitly demarcates the goal directedness of the fourth order from apparent goals implied in the teleonomic operation of living things implicit in the third. We must therefore assume he means active goal-setting : the exercise of what we commonly refer to as agency or free will. Agency results from the vastly expanded behavioral plasticity available when an organism develops an advanced nervous system. Also, to learn an agent must have some form of memory. Memory too is generally associated with the existence of a central nervous system and is often seen as involving stored representations. But the idea of 'representations' is highly problematic from a biological point of view. What is it that is represented and how? We consider this problem in the next section.

Ellis would seem to be pointing to a category here which deals with non-human animals but the transition points are not well defined from the perspective of mechanisms of emergence. Learning in animals can stretch from simple operant conditioning to complex evaluative processes involving logical reflexion. Different stages along this continuum would appear to support significantly different forms of emergent structure. Ellis makes no distinction, for example, between individual and social learning.

Ellis marks his final transition from category four to category five by moving from simple learning capability to the capacity for language. Animals such as apes have rudimentary language ability – are they included in here or is this category the human catch-all category? Unfortunately the more closely we look at the jump between fourth and fifth order the more it resembles an abyss.

There has been a considerable research effort directed at understanding the origins and developmental phases associated with the attainment of the distinctive human cognitive capabilities. These are the capabilities which seem to relate to the transition between Ellis' category four and five orders of emergence.. Much of this has drawn on comparative neurology, and sociological and psychological study of non-human animals, in particular apes. Insights are available also from developmental psychology and neurology directed at understanding Human ontogeny: the phases of development from infant to adult. Note that these may overlap as phylogenetically determined capabilities characteristic of some animals may correspond to early stages of human ontogenetic development. This corpus offers those of is involved with AI two opportunities a) a capacity to aim to better stage the development of agent specifications - aiming to provide a reasonable model for simple intelligence before the more complex and b) a capacity, even before we can effectively model or simulate more advanced intelligence, to theorize about the implications it may have for emergence of social structure.

Some work has already been undertaken in this area, most notably in the area of robotics rather than computer simulation of social phenomena (although robots can be regarded as physical simulations and multi-agent software simulations as simulated robotics). Of particular note here is the work of Dautenhahn (2001; 2002), Bryson (2007; n.d) and Steels (1997; 2005; 1999) in the area of language.

Gardenfors (2006) identifies the following as needing to be explained (presented in order of their apparent evolution).

- Sensations
- Attention
- Emotions
- Memory
- Thought and imagination
- Planning
- Self-consciousness/theory of mind
- Free-will
- Language

These are present to varying degrees in different organisms and develop at different stages in humans as they develop from infancy to adulthood. The degree of interrelatedness is not, however, straight forward. Apes for example demonstrate self-awareness and theory of mind but do both without language whereas in humans language appears to play a significant role in both. For the time being then too little is known about these transitions.

It is perhaps in understanding these transitions that we find the greatest challenges for advancing artificial societies and it is here that we find philosophy may have dealt us an unhelpful turn. The advent of the central nervous system and the observation that cognitive function is correlated with brain size has contributed to a distinctive account of the function of brain and its relationship to mind (Johnson, 1990; Lakoff & Johnson, 1999). In this convention, mind and hence cognition has been argued to originate in brains and to involve symbol manipulation. As we consider the literature on what makes human cognition distinctive, we need to be mindful of the effect of this and alternative paradigms. What are these alternatives and what difference do they make to our understanding of orders of emergence in general and social emergence in particular?

TWO PARADIGMS – TWO POSSIBLE APPROACHES

Within AI there are two alternative and some argue antithetical paradigms of cognition – symbolic and connectionist. Symbolic AI assumes that it is possible to model every general intelligence using a suitable symbol system and that intelligence involves symbol manipulation (Franklin, 1998).

In their *The Embodied Mind*, Varela & Rosch (1992) state:

> The central intuition … is that intelligence—human intelligence included—so resembles computation in its essential characteristics that cognition can actually be defined as computations of symbolic representations (F. Varela, Thompson, & Rosch, 1992: 40).

The symbolic approach inevitably constructs a duality. The environment is experienced as a facticity and acted upon directly, but is also conceived and symbolically represented in the mind. Mind and behaviour are linked as hypothesis and experiment. The mind looks for patterns in representations and tests the degree to which these accord with the outside world.

More recently, this tradition has been challenged. The advent of complexity theory has given greater impetus to connectionist models of mind such as neural networks. Here emergent structure or pattern arises from massively interconnected webs of active agents. Applied to the brain, Varela et al state:

> The brain is thus a highly cooperative system: the dense interconnections amongst its components entail that eventually everything going on will be a function of what all the other components are doing (1992: 94).

It is important to note that no symbols are invoked or required by this model. Meaning is embodied in fine-grained structure and pattern throughout the network. Unlike symbolic systems, connectionist approaches can derive pattern and meaning by mapping a referent situation in many different (and context dependent) ways. Meaning in connectionist models is embodied by the overall state of the system in its context. It is implicit in the overall 'performance in some domain'. Herein lays its major problem from the perspective of multi-agent simulation. In connectionist models the micro-states which support a given macro state is opaque – relatively inaccessible to an observer and difficult to interpret – indeed, there will often be several or many micro configurations compatible with a given macro-state (Richardson, 2002b). Several attempts have been made to address this problem. The first was to consider hybrid systems in an attempt to gain the advantage of each (Khosla & Dillon, 1998). The second has been to find a middle ground. This is apparent for example in Gardenfors' theory of conceptual spaces (Gardenfors, 2004). At the same time the practical value of connectionist systems – their capacity to categorize contexts or situations in a non-brittle way– has been seen as a significant advantage in robotics (Brooks, 1991).

Back in 1992 Varela et al noted that:

> …an important and pervasive shift is beginning to take place in cognitive science under the very influence of its own research. This shift requires that we move away from the idea of the world as independent and extrinsic to the idea of a world as inseparable from the structure of [mental] processes of self modification. This change in stance does not express a mere philosophical preference; it reflects the necessity of understanding cognitive systems not on the basis of their input and output relationships but by their *operational closure* (1992: 139).

They go on to argue that connectionist approaches, while an advance on cognitivism are not consistent with an approach which views biological agents as operationally closed in that *'…the results of its processes are those processes themselves'* (1992, p. 139). They assert:

> Such systems do not operate by representation. Instead of *representing* an independent world, they *enact* a world as a domain of distinctions that is inseparable from the structure embodied by the cognitive system (1992: 140).

These authors argue for an approach of cognition as 'enaction', an intertwining of experience and conceptualization which results from the structural coupling of an autonomous organism and its environment. Autopoietic theory provided a concrete and operationalizable account of the intertwining of micro and macro at the level of the cell. The enactive theory of cognition goes some way towards providing a basis for understanding this process in multi-cellular animals. Enactive cognition is currently enjoying significant attention and hence conceptual extension as well as experimental grounding in the field of robotics (see for example De Jaegher & Di Paolo, 2007; Di Paolo et al., 2007; Metta, Vernon, & Sandini, 2005). The attraction here is pragmatic – it helps to address longstanding problems within robotics, in particular the problem of symbol grounding (Harnad, 1990). To date it has seen little uptake within social simulation. The implications of enaction go well beyond pragmatics however.

The enactive turn in AI has as an explicit target a resolution of the micro-macro problem. While symbolic AI assumes the existence of an objective independent world and a mental model with some correspondence to the real world, enaction dispenses with this dichotomy. As an autonomous entity, the cognizing agent is concerned only to maintain its viability in an environment. It adjusts its structure to accommodate perturbation from the environment (which includes other cognitive agents) in order to do so. Advanced nervous systems and capabilities such as language simply extend the requisite variety available to the agent extending the range and type of environmental perturbations it can survive. As agents and environments structurally couple they co-determine one another to 'satisfice' the conditions for mutual viability. From this perspective, the importance of environment recedes from determinant to constraint. Intelligence moves from problem solving capacity to flexibility to enter into and engage with a shared world. However, McGee (2005a; 2005b) has recently argued that despite its promise, enactive cognition is not yet sufficiently well articulated to 'speak of hypothetical mechanisms'. The limiting factor here would appear to be as much one of insufficient application as theoretical difficulty. In the final section we attempt a definition of two classes of emergence which we call reflexive and non-reflexive. These draw on the enactive paradigm and attempt to provide a concrete specification of the mechanisms which underlay each.

TOWARDS AN ENACTIVE SPECIFICATION OF ASPECTS OF COGNITION AND THEIR ASSOCIATED ORDERS OF EMERGENCE

How then do we advance our understanding of the effect f different cognitive capability on orders of emergence? A useful strategy may be to simplify the problem. By way of a mental exercise we will take simple extremes and recast the problem in terms of an enactive view. From an enactive position the critical phases of cognitive development appear to be as follows:

- Autonomy (operational closure)
- Structural Coupling
- Reflexivity/self consciousness

- Language/consensual domains

All living beings (from amoeba to humans) are distinguished by autonomy and as autonomous entities they necessarily enter into structural coupling with their environment. We take this as one pole of the continuum and identify the class of emergence which it can support as non-reflexive. This is the enactive equivalent to social order which is a product of emergence *without* the feedback loop from macro to micro which Castelfranchi (1998a) refers to as immergence. The mechanisms are, however, more sophisticated than are currently modeled in Artificial Societies as they involve autonomous agents – these are essentially what Ellis refers to in his category four – i.e. biological agents which can change their structure (learn) in response to environmental perturbation. It should be feasible to simulate this type of agent with current technology or at least to achieve a close proxy although we have not yet managed to do so beyond the most basic chemical system analogues of cell autopoiesis. If we were to achieve it how might we describe the system operation?

Non-reflexive social emergence

Non-reflexive emergence arises from the mechanism of structural coupling between operationally closed (autonomous) agents. Structural coupling will arise between such agents which have sufficient cognitive range (behavioral repertoire) when they are located in a common environment. Assuming that their phylogeny and ontogeny is such that they can co-exist, through the process of recurrent mutual perturbation, each will adjust its structure so as to accommodate the other – their structures will become mutually aligned or structurally coupled. This process has been approximated in a simulation by Stoica-Kluver and Kluver (2006).

An observer may notice regularities in the resulting patterns of interaction and these may be labeled as 'norms' for example although Castelfranchi would refer to them as social functions as they 'work without being understood'. These patterns represent mutual accommodations, and an observer might attribute to those accommodations some social 'function'. The accommodations an agent makes to remain viable in one domain of interaction will need to be reconciled (within its body-hood) against accommodations being made (simultaneously) as it also participates with different agents in other domain/s in which it is simultaneously participating – agency and structure converge and are both instantiated at the point of enaction. The accommodations made will be those that allow the agent to remain viable and to maintain its organization (i.e. which 'satisfice' the constraints and allow conservation of identity) based on its unique ontogeny (structure resulting from its history of interactions in a variety of domains including the current one).

Here the emergent structure can be seen to be 'in' (i.e. internalized within its own cognitive structure) each agent to the extent that each has had to make structural adjustments to operate in the shared domain. The structural adjustment each needs to make in order to persist will, however, be unique. In other words the structural accommodations each has made in order to contribute to the patterns, will *not* be the same. The structure, then, can also be regarded as 'in' the network, as it is the intersection of these disparate agent structures which gives it its particular form at a particular time. As any agent could leave the domain and have minimal effect on the resulting pattern, each agent's 'contribution' will be relatively small. The pattern can be thought about as like a hologram. The whole is in every part (agent) such that removal of parts (agents) reduces the resolution (coherence) but does not constitute loss of overall pattern. However, the loss of too many components may reduce the coupling to the point that the existing pattern de-coheres and transforms into something different. Each agent contributes to the pattern formation, so it is conceivable that the pattern will only be realized with some critical minimal number of agents present which have had a sufficient mutual history to have aligned their structures.

In natural systems, the local level interactions between agents are constrained by the existing structures of the agents and the state of their environment. With biological agents the system is open in that any emergent structure is possible as long as it remains consistent with the biological viability of the agents as living (autopoietic) entities. This biological constraint includes limits to environmental conditions conducive to life (i.e. not too hot or too cold, the need for energy, limitations to sensory channels, channel bandwidths and affective/psychomotor response capabilities etc). These are primarily a product of phylogeny (the evolutionary history of the organism at the level of the species) rather than ontogeny (the history of development at the level of the individual), and are therefore slow to change and not under the control of the emergent social system. As a consequence the basic dimensionality of the phase space of the social system does not change over the time frame of interest for understanding social systems. The dimensionality of the phase space is determined by the dimensions of variability possible by individuals – i.e. the plasticity of their nervous systems and by higher order dimensions which emerge from their interaction.

Reflexive Social Emergence

What changes if we now jump to the opposite pole on our hypothetical continuum? Here we attempt to outline the difference made by agents which are self aware and which can interact in language.

Biological agent's sensory surfaces are selected to be sensitive to difference in dimension of their world relevant to their survival and their cognitive apparatus is thus geared to make distinctions relevant to maintaining their viability in past environments. Once cognitive complexity exceeds a critical threshold (Gardenfors, 2006) these distinctions can be represented in language. Maturana and Varela (1980) describe language as involving the co-ordination of the co-ordination of actions – i.e. language provides a meta process by which agents orientate themselves within a world. Structural coupling can arise purely through behavioral coordination of action (as discussed above), but it can also take place in and through linguistic exchange – the mutual co-ordination of co-ordination of behaviors. This gives rise to a consensual linguistic domain characterized by a more or less shared lexicon. This process has been simulated using both shared referents and simple structural coupling in the absence of objective referents (Gong, Ke, Minett, & Wang, 2004; Hutchins & Hazlehurst, 1995; Steels, 1997, 1998; Steels, 2005; Steels & Kaplan, 1998; Steels & Kaplan, 1999), as has the emergence of a rudimentary grammar (Howell & Becker, n.d; Vogt, n.d).

The advent of language radically increases the behavioral plasticity of agents and has significant implications for the dimensionality of the phase space and of the resulting higher order structures it can generate and support. This is because language makes possible the emergence of domains of interaction which can themselves become the target for further linguistic distinction and hence new domains. In other words, language allows the agent to make distinctions on prior distinctions (to language about its prior language or to build further abstractions on prior abstractions). This supports the possibility of infinite recursion and infinite branching (there are no doubt biological constraints on this in humans). This is an intrinsically social process. Furthermore, a capacity to distinguish (label or categorize) processes supports reification and this simplifies the cognitive handling of processual phenomena and allows the resulting reifications to be treated by the agent in the same manner as material objects.

These capabilities greatly expand the structural flexibility of the agents: they can now invent shared epistemic worlds. The phase space of agent cognition is now based primarily on constraints of ontogeny rather than phylogeny and is hence under the influence of the agent/s.

Language makes possible a further major qualitative difference in natural and human social emergence. Humans (and possibly some other primates, cetaceans and elephants)[2] have developed sufficient cognitive capacity to become self-aware and as such exhibit reflexive behavior. This occurs when the agent is capable of distinguishing 'self' and 'other' i.e. the agent can entertain the notion of 'I' as a concept and treat that concept as an object. The advent of this capacity for reflexive identity also supposes the existence of a range of conceptual operators that act on identity – identity construction and maintenance becomes a part of the agent's world creation. Exploration of this process is proceeding under the title of Neurophenomenology (Rudrauf, Lutz , Cosmelli, Lachaux , & Le Van Quyen, 2003; Thompson & Varela, 2001).

In other words, agents can now notice the patterns that arise as they interact with others and distinguish those patterns in language. Such a mechanism would be the enactive equivalent to Castelfranchi's (1998a) Cognitive Emergence. Here a reflexive agent can notice an emergent pattern of social behavior and explicitly denote it as a 'norm' for example. While this denotation may be idiosyncratic (i.e. based on the necessarily limited perception of the individual agent), the agent can nonetheless act on the basis of this denotation. Once distinguished and reified within a domain, agents can decide (on the basis of rational as well as value based or emotional criteria) how to respond – they can choose to ignore the norm or to behave in ways they believe will limit the reoccurrence of the behaviors that are outside the agreed/shared patterns of the group. Once a pattern has been distinguished in language it can make the transition to a rule: a formally stated, linguistically explicit requirement with stated conditionals and possible resources to maintain it. This suggests that an agent can form hypotheses about the relationship between a macro structural aspect of the social system in which it is a participant and then act on that hypothesis, potentially changing the structure which it participates in generating. This gives rise to a feedback path between macro and micro phenomena that is not present in any other natural phenomena.

Consistent with Castelfranchi's claim, agents possessing this cognitive complexity form the components of a social system which would exhibit a distinct class of emergence. From the emergent perspective this is argued on the basis that reflexive agents will display qualitatively different behaviors from non-reflexive through the ability to modify their own sets of behavioral change triggers. For agents which have linguistic capability, the two processes (linguistic and non-linguistic) intertwine or even become one and would not be able to be empirically disentangled. Their respective influences will only be able to be examined through simulations or by comparing agents with different (phylogenetic) capabilities (i.e. different species) and this sets some interesting methodological challenges.

The Role of the Observer

Another significant implication of the relationships described above is the observer dependant nature of emergence in social systems. In human social systems every agent is an observer and it is the process of observation and the associated distinction-making which is the reflexive engine of emergence. In natural systems, the agents of the system are unable to observe and distinguish linguistically or to distinguish external structures as separate from

[2] It is important to note that we can infer the existence of threshold effects here but cannot precisely specify the critical points of complexity at which self-awareness and language becomes possible. The ability for language is of course evident in species other than humans, but the degree to which their linguistic plasticity involves or enables reflexivity in the system is a subject for further research.

themselves hence the process of observation has no impact on the dynamics of the system or the way in which emergence takes place. To some extent we can see an acknowledgement of this effect in methodological discussions within ethnography, action research (Carr & Kemmis, 1986) and grounded theory (Corbin & Strauss, 1990). In each of these methodologies the impact of the researcher on the social system under study is acknowledged and seen as part of the process. The view being proposed here is that any agent that becomes a part of the system being observed has the potential to influence that system. An agent can become a part of the system simply by being itself observed or conceived as observing by those who constitute the system. In other words, the effect of the entry of a new observing agent is to change the system boundary so as to include that agent. The boundary is itself an entity of ambiguous status – it is an epistemic distinction albeit one based on potentially ontological markers. In most social theory, positing the observer as a necessary part of the system removes any ontological privilege and threatens either infinite recursion or paradox. Based on the position advocated here, a degree of both may well be fundamental to the type of system being described (Hofstadter, 2007).

Implications for emergence

Complex systems of all kinds demonstrate a capacity to give rise to complex macro patterns as a result of local interactions between agents in highly connected webs. This local interaction can often be characterized as involving some signaling between agents. As we have seen above, in human social systems, this signaling behavior takes on a qualitatively different form. This has three key implications for our understanding of emergence that to date have largely been ignored by the literature.

1. *Social systems will display an increased range of emergent possibilities*: the reflexive nature of social systems implies that a greater range of emergent structures should be expected and they will be subject to more rapid change.
2. *Dimensions of phase space are non-constant*: As the agents in the social system define and redefine the phase space as a function of their reflexive distinctions they will create and change the dimensions of that phase space, in order to support their own viability in that space.
3. *Phase space comes under control of the system and is dynamic*: The dimensionality of the phase space associated with ontogenetic parameters is derived through the self-distinguishing characteristics of the agents and can be influenced by their situated behavior. Significantly the feedback path between macro and micro would add significant non-linearity to the system and it becomes important to identify and explain order producing mechanisms within the network.

CONCLUSION AND FUTURE DIRECTIONS

In this chapter we have attempted to provide an operational specification of the gap implicit in Ellis' fourth and fifth order emergence. In a sense we have demarcated the extremes using the lens of enactive cognition. Enactive cognition was selected as it provides a theoretical underpinning which avoids the dualism inherent in symbolic systems and the confusion of fundamental processes which results from this. It has been argued to be both theoretically better capable of capturing the essential mechanisms and of providing a practical way of avoiding the now well documented pitfalls of symbolic AI. From this perspective the first challenge that must be addressed to advance social simulation is to achieve some form or proxy of constitutive autonomy in our multi-agent models. Significant work is currently underway on this problem in robotics but there have been few systematic attempts within social simulation.

Once this has been achieved we then need to model autonomous closure in linguistic systems. We would seem to be a very long way from this at present. It may be possible however to achieve this first in some abstract domain – simulating perhaps Luhmann's self-referential systems of communicative acts. This is probably unlikely however.

In our sketching out the extremes many questions remain about what might lay in the middle. This middle includes very significant phases of human cognitive development – including theory of mind and narrative intelligence. There can be no doubt that these will support qualitatively distinct classes of emergent social phenomena. There is evidence from the study of apes that forms of these cognitive capabilities do not require language. These may be much more accessible to our still limited capacity to simulate than the human equivalents which appear to intertwine with linguistic capability. We probably have much to learn then from the study of primate communities and from research into cognition in species other than humans. At present these attract considerable less attention within the social simulation community and perhaps this is a mistake. We have learned a lot from ants – how much more from apes? Robotics also appears well equipped to incorporate the insights coming from situated, embodied and enactive cognition. It s more difficult to see how embodied proxies may be incorporated into multi-agent simulations but no doubt there are ways. Such systems will doubtless need to be able to bootstrap some level of operational closure and it will be behavior within the self-determining boundary that – free from the inevitable teleological hand of the designer can reveal insights into how we humans do what we seem to do so effortlessly – construct social worlds in which we can live viable and interesting lives.

REFERENCES

Ablowitz, R. (1939). The Theory of Emergence. *Philosophy of Science, 6*(1), 16.

Archer, M. (1998). Realism in the Social Sciences. In M. Archer, R. Bhaskar, A. Collier, T. Lawson & A. Norrie (Eds.), *Critical Realism: Essential Readings*. London: Routledge.

Archer, M., Bhaskar, R., Ciollier, A., Lawson, T., & Norrie, A. (1998). *Critical Realism: Essential Readings*. London: Routledge.

Axelrod, R. (1984). *The Evolution of Cooperation*. New York: Basic Books.

Bak, P. (1996). *How Nature Works: The Science of Self-Organized criticality*. New York: Copurnicus.

Barandiaran, X. (2005). Behavioral Adaptive Autonomy. A Milestone on the ALife route to AI? San-sebastian, Spain: Department of Logic and Philosophy of Science, University of the Basque Country.

Bednarz, J. (1988). Autopoesis: the Organizational Closure of Social Systems. *Systems Research, 5*(1), 57-64.

Berger, P. L., & Luckman, T. (1972). *The Social Construction of Reality*: Penguin.

Bertalanffy, L. v. (1950). An Outline of General Systems Theory. *British Journal for the Philosophy of Science, 1*(2).

Bertalanffy_von, L. (1968). *General Systems Theory*. New York: Braziller.

Bhaskar, R. (1997). A Realist Theory of Science. In. London: Verso.

Bhaskar, R. (1998). *The Possibility of Naturalism*. London: Routledge.

Brooks, R. A. (1991). Intelligence without representation. *Intelligence without reason*(47), 569-595.

Bryson, J. J. (2007). Embodiment vs Memetics. Bath: Artificial Models of Natural Intelligence, University of Bath.

Bryson, J. J. (n.d). Representational Requirements for Evolving Cultural Evolution, *Interdiciplines*.

Carr, W., & Kemmis, S. (1986). *Becoming Critical: Knowing Through Action Research*: Deakin University.

Castelfranchi, C. (1998a). Simulating with Cognitive Agents: The Importance of cognitive emergence. In J. S. Sichman, R. Conte & N. Gilbert (Eds.), *Multi-agent Systems and Agent Based SImulation*. Berlin: Springer.

Castelfranchi, C. (1998b). Simulating with Cognitive Agents: The Importance of Cognitive Emergence. In J. S. Sichman, R. Conte & N. Gilbert (Eds.), *Lecture Notes in Artificial Intelligence*. Berlin: Springer Verlag.

Chalmers, D. J. (2006). Strong and Weak Emergence. Canberra: Research School of Social Sciences, Australian National University.

Checkland, P. (1988). *Systems Thinking Systems Practice*. G.B.: John Wiley.

Clayton, P. (2006). Conceptual Foundations of Emergence Theory. In P. Clayton & P. Davies (Eds.), *The re-Emergecne of Emergence: The Emergentist Hypothesis from Science to Religion*. Oxford: Oxford University Press.

Clayton, P., & Davies, P. (2006). *The Re-Emergence of Emergence: The Emergentist Hypothesis from Science to Religion*. Oxford: Oxford University Press.

Coleman, J. S. (1994). *Foundations of Social Theory*. Cambridge: Belknap.

Conte, R., Hegselmann, R., & Terna, P. (1997). *Simulating Social Phenomena*. Berlin: Springer.

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology, 13*(1), 18.

Dautenhahn, K. (2001). The Narrative Intelligence Hypothesis: In Search of the Transactional Format of Narratives in Humans and Other Social Animals. In *Cognitive Technology* (pp. 248-266). Hiedelberg: Springer-Verlag.

Dautenhahn, K. (2002). The Origins of Narrative. *International Journal of Cognition and Technology, 1*(1), 97-123.

Davies, P. (2006). The Physics of Downward Causation. In P. Clayton & P. Davies (Eds.), *The Re-Emergence of Emergence: The Emergentist Hypothesis from Science to Religion*. Oxford: Oxford University Press.

De Jaegher, H., & Di Paolo, E. A. (2007). Participatory Sense-making: An enactive approach to Social Cognition. *Phenomenology and thec ognitive Sciences, forthcoming*.

Di Paolo, E. A., & Lizuka, H. (2007). How (not) to Model Autonomous Behaviour. *Biosystems*.

Di Paolo, E. A., Rohde, M., & De Jaegher, H. (2007). Horizons for The Enactive Mind: Values, Social Interaction and Play. In J. Stewart, O. Gapenne & E. A. Di Paolo (Eds.), *Enaction: Towards a New Paradigm for Cognitive Science*. Cambridge MA: MIT Press.

Ellis, G. F. R. (2006). On the Nature of Emergent Reality. In P. Clayton & P. Davies (Eds.), *The Re-Emergence of Emergence: The Emergentist Hypothesis from Science to Religion*. Oxford: Oxford University Press.

Engels, F. (1934). *Dialectics of Nature*. Moscow: Progress Publishers.

Epstein, J. M., & Axtel, R. (1996). *Growing Artificial Societies*. Cambridge, Ma.: MIT Press.

Eronen, M. (2004). *Emergence in the Philosophy of Mind.* University of Helsinki, Helsinki.

Fodor, J. A. (1974). Special; Sciences or The Disunity of Science as a Working Hypothesis. *Synthese, 28*, 18.

Franklin, S. (1998). *Artificial Minds*. London: MIT press.

Froesea, T., Virgo, N., & Izquierdo, E. (2007). Autonomy: a review and a reappraisal. Brighton Uk: University of Sussex.

Fuchs, C., & Hofkirchner, W. (2005). The Dialectic of Bottom-up and Top-down Emergence in Social Systems. *tripleC 1*(1), 22.

Gardenfors, P. (2004). *Conceptual Spaces*. London: The MIT Press.

Gardenfors, P. (2006). *How Homo became Sapiens: On the evolution of Thinking*. Oxford: Oxford University Press.

Gell-Mann, M. (1995). *The Quark and the Jaguar: Adventures in the simple and the complex*. Great Britain: Abacus.

Giddens, A. (1984). *The Constitution of society: Outline of the theory of structuration*. Berkeley: University of California Press.

Gilbert, N. (1995). Emergence in Social Simulation. In N. Gilbert & R. Conte (Eds.), *Artificial Societies*. London: UCL Press.

Gilbert, N. (2002). *Varieties of Emergence.* Paper presented at the Social Agents: Ecology, Exchange, and Evolution Conference Chicago.

Goldspink, C. (2000). *Social Attractors: An Examination of the Applicability of Complexity theory to Social and Organisational Analysis.* Unpublished PhD, University Western Sydney, Richmond.

Goldspink, C., & Kay, R. (2003). Organizations as Self Organizing and Sustaining Systems: A Complex and Autopoietic Systems Perspective. *International Journal General Systems, 32*(5), 459-474.

Goldspink, C., & Kay, R. (2004). Bridging the Micro-Macro Divide: a new basis for social science. *Human Relations, 57* (5), 597-618.

Gong, T., Ke, J., Minett, J. W., & Wang, W. S. (2004). A Computational Framework to Simulate the co-evolution of language and social structure.

Harnad, S. (1990). The Symbol Grounding Problem. *Physica, 42*, 335-346.

Hofstadter, D. R. (2007). I am a Strange Loop. In: Basic Books.

Holland, J. H. (1998). *Emergence: from chaos to order*. Ma.: Addison Wesley.

Howell, S. R., & Becker, S. (n.d). Modelling Language Aquisition: Grammar from the Lexicon?

Hutchins, E., & Hazlehurst, B. (1995). How to invent a lexicon: the development of shared symbols. In N. Gilbert & R. Conte (Eds.), *Artificial Societies*. London: UCL Press.

Jackson, M. C. (2000). *Systems Approaches to Management*. London: Kluwer Academic.

Johnson, M. (1990). *The Body in the Mind: The Bodily Basis of Meaning, Imagination and Reason*. Chicago: The University of Chicago Press.

Kauffman, S. (2000). *Investigations*. New York: Oxford.

Kauffman, S. A. (1993). *The Origins of Order: Self Organization and Selection in Evolution*: Oxford University Press.

Kauffman, S. A. (1996). *At home in the Universe: The Search for Laws of Complexity*. London: Penguin.

Kay, R. (1999). *Towards an autopoietic perspective on knowledge and organisation.* Unpublished PhD, University of Western Sydney, Richmond.

Keeney, B. P. (1987). *Aesthetics of Change*: Guilford.

Kennedy, J., & Eberhart, R. C. (2001). *Swarm Intelligence* (1 ed.). London: Academic Press.

Khosla, R., & Dillon, T. S. (1998). Welding Symbolic AI with Neural Networks and their APPLICATIONS. *IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION*.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western thought*. New York: Basic Books.

Leont'ev, A. N. (1978). *Activity, Consciousness and Personality*. Engelwood Cliffs: Prentice Hall.

Lorenz, E. N. (2001). *The Essence of Chaos* (4 ed.). Seattle: University of Washington Press.

Luhmann, N. (1990). *Essays on Self Reference*. New York: Columbia University Press.

Luhmann, N. (1995). *Social Systems*. Stanford: Stanford University Press.

Maturana, H., & Varela, F. (1980). *Autopoiesis and Cognition: The Realization of the Living* (Vol. 42). Boston: D. Reidel.

Maturana, H. R., & Varela, F. J. (1992). *The Tree of Knowledge: The Biological Roots of Human Understanding*. Boston: Shambhala.

McGee, K. (2005a). Enactive Cognitive Science. Part 1: Background and Research Themes. *Constructivist Foundations, 1*(1), 15.

McGee, K. (2005b). Enactive Cognitive Science. Part 2: Methods, Insights, and Potential. *Constructivist Foundations, 1*(2), 9.

McKelvey. (1997). Quasi-Natural Organisation Science. *Organization Science, 8*, 351-380.

McMullin, B., & Grob, D. (2001). Towards the Implementation of Evolving Autopoietic Artificial Agents, *6th European Conference on Artificial Life ECAL 2001*. University of Economics, Prague.

Metta, G., Vernon, D., & Sandini, G. (2005). *The Robotcup Approach to the Development of Cognition.* Paper presented at the Fifth International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems Lund University Cognitive Studies, .

Mingers, J. (2002). Are Social Systems Autopoietic? Assessing Luhmanns Social Theory. *Sociological review, 50*(2).

Mingers, J. (2004). Can Social Systems be Autopoietic? Bhaskar's and Giddens' Social Theories. *Journal for the Theory of Social Behaviour, 34*(4), 25.

Moreno, A., & Etxeberria, A. (1995). Agency in natural and artificial systems. San Sabastian, Spain: Department of Logic and Philosophy of Science University of the Basque Country.

Moreno, A., Umerez, J., & Ibanes, J. (1997). Cognition and Life. *Brain and Cognition, 34*, 107-129.

Oyama, S. (2000). *The Ontogeny of Information: Developmental Systems and Evolution*: Duke University Press.

Peterson, G. R. (2006). Species of Emergence. *Zygon, 41*(3), 22.

Prigogine, I. (1997). *The End of Certainty: Time, Chaos and the New Laws of Nature*. New York: The Free Press.

Prigogine, I., & Stengers, I. (1985). *Order out of Chaos: Man's New Dialogue with Nature*: Flamingo.

Richardson, K. A. (2002a). Methodological Implications of a Complex Systems Approach to Sociality: Some further remarks. *Journal of Artificial Societies and Social Simulation, 5*(2).

Richardson, K. A. (2002b). *On the Limits of Bottom Up Computer Simulation: Towards a Non-linear Modeling Culture.* Paper presented at the 36th Hawaii International Conference on Systems Science, Hawaii.

Rocha, L. M. (1998). Selected Self-Organization: and the semiotics of evolutionary systems In S. Salthe, G. Van de Vijver & M. Delpos (Eds.), *Evolutionary Systems: Biological and Epistemological Perspectives on Selection and Self-Organization* (pp. 341-358): Kluwer Academic Publishers.

Rudrauf, D., Lutz , A., Cosmelli, D., Lachaux , J.-P., & Le Van Quyen, M. (2003). From Autopoiesis to Neurophenomenology: Francisco Varela's exploration of the biophysics of being. *Biol. Res, 36*, 27-65.

Sawyer, K. R. (2001). Emergence in Sociology: Contemporary Philosophy of Mind and Some Implications for Sociology Theory. *American Journal of Sociology, 107*(3), 551-585.

Sawyer, K. R. (2003). Artificial Societies: Multiagent Systems and the Micro-macro Link in Sociological Theory. *Sociological Methods & Research, 31*, 38.

Shrader, W. E. (2005). *The Metapysics of Ontological Emergence.* University of Notre Dame.

Stanford Encyclopedia of Philosophy. (2006). Emergent Properties, *Stanford Encyclopedia of Philosophy*.

Steels, L. (1997). *Constructing and Sharing Perceptual Distinctions.* Paper presented at the European Conference on Machine Learning, Berlin.

Steels, L. (1998). *Structural coupling of cognitive memories through adaptive language games.* Paper presented at the The fifth international conference on simulation of adaptive behavior on From animals to animats 5, Univ. of Zurich, Zurich, Switzerland.

Steels, L. (2005). The emergence and evolution of linguistic structure: from lexical to grammatical communication systems. *Connection Science, 17*(3 & 4), 17.

Steels, L., & Kaplan, F. (1998). Stochasticity as a Source of Innovation in Kanguage Games. In C. Adami, R. K. Belew, H. Kitano & C. Taylor (Eds.), *Artificial Life VI*. Cambridge, MA: MIT Press.

Steels, L., & Kaplan, F. (1999). Bootstrapping grounded word semantics. In T. Briscoe (Ed.), *Linguistic evolution through language acquisition: formal and computational models,* . Cambridge, UK: Cambridge University Press.

Stewart, I. (1990). *Does God Play Dice - The New Mathematics of Chaos*: Penguin.

Stioica-Kluver, C., & Kluver, J. (2006). Interacting Neural Networks and ther Emergence of Social Structure. *Complexity, 12*(3), 11.

Thompson, E., & Varela, F. J. (2001). Radical Embodiment: neural dynamics and consciousness. *TRENDS in Cognitive Sciences, 5*(10), 418-425.

Varela, F. (1979). *Principles of Biological Autonomy*. New York: Elsevier-North Holland.

Varela, F. (1997). Patterns of Life: Intertwining Identity and Cognition. *Brain and Cognition, 34*, 72-87.

Varela, F., Maturana, H., & Uribe, R. (1974). Autopoiesis: The Organization of Living Systems, Its Characterization and a Model. *Biosystems, 5*, 187-196.

Varela, F., Thompson, E., & Rosch, E. (1992). *The Embodied Mind*. Cambridge: MIT Press.

Vogt, P. (n.d). Group Size Effects on the Emergence of Compositional Structures in Language. Tilburg, Netherlands: Tilburg University.

von_Krogh, G., & Roos, J. (1995). *Organizational Epistemology*. London: St Martins Press.

Vygotsky, L. S. (1962). *Thought and Language*. Cambridge, Mass: MIT Press.

Williams, G. P. (1997). *Chaos Theory Tamed*. Washington D.C: Joseph Henry Press.

Zeleny, M. (1991). *Autopoiesis: A Theory of Living Organization*. New York: North Holland.